10-24-2018

Slides for workshop 'On-Campus Resources for Federal Grant Seekers'

David Luftig
University of Dayton, dluftig1@udayton.edu

DATA MANAGEMENT RESOURCES FOR FEDERAL GRANT SEEKERS

David Luftig

Research Services Librarian

dluftig1@udayton.edu

University of Dayton

Campus Conversation

Learning and Teaching Center

October 24th, 2018

Roesch Library Research Services

Can assist with…

- Data management and data management plans (DMPs)
- Creating and linking ORCIDs and Researcher IDs
- Finding data sets, literature, patents, standards, etc.
- Understanding the research and grant landscape through research metrics
- Evaluating data repositories and open access publishing
- Evaluating journals (through the use of metrics) and issues of licensing

FEDERAL DATA REQUIREMENTS

NSF Final Publication Requirements

NSF requires that either the version of record or **the final accepted manuscript** in peer-reviewed scholarly journals and papers in juried conference proceedings or transactions **be deposited in a public access compliant repository** designated by NSF; **be available for download, reading and analysis** free of charge **no later than 12 months after initial publication**; possess a minimum set of machine-readable metadata elements in a **metadata record** to be made **available free of charge** upon initial publication; be managed to **ensure long-term preservation**; and be reported in annual and final reports during the period of the award with a persistent identifier that provides links to the full text of the publication as well as other metadata elements. https://www.nsf.gov/pubs/2016/nsf16009/nsf16009.jsp#q1

NSF Data Set Requirements

Investigators are **expected to share** with other researchers**, at no more than incremental cost** and **within a reasonable time**, the **primary data, samples, physical collections and other supporting materials** created or gathered in the course of work under NSF grants….Grantees are expected to encourage and facilitate such sharing. Privileged or **confidential information should be released only in a form that protects the privacy** of individuals and subjects involved.
https://www.nsf.gov/pubs/policydocs/pappg17_1/pappg_11.jsp#XID4

In General, DMPs…

- Have been required since 2011 for federal research grants
- Can have slightly different looks based on the discipline of the research (such as engineering or earth sciences) as well as the granting agency (such as the NIH or the NSF)
- Are no longer than two pages and are standardized
- Provide a clear outline of how you're going to collect, store, share, label, and preserve your data
- Identify who the PIs are and the data collection roles of the researchers working on the project

Benefits of Creating a Data Management Plan

- All project researchers are aware of expectations, standards, and responsibilities prior to, and during, data collection
- Data can be located and utilized at the time when it is needed
- Data and other documentation is less likely to be missing (when it comes time to process or share results)
- There is continuity in the process, even if new researchers join the project
- Avoids unnecessary duplication and redundancy
- Data can be easily understood so as to allow for quicker publishing, sharing of results, and data validation

Considerations of Research Data (and a DMP)

Data Collection

What types of data will you be utilizing? How will you collect and use the data? Is there special equipment used to collect the data? What is the medium of the data (digital, analog, etc.)? Who will be responsible for the data?

Data Description

What file formats and naming conventions will be used? How will data be organized? How will data be documented? How will data be described?

Storage

Where and how will you store your data? What is your backup plan? Who has access to your data? How will you manage security?

Protection, Privacy, Ownership

Are there ethical or privacy concerns? Who owns the data? Are there restrictions on the data? What will happen to the data if the PI leaves the institution?

Sharing and Access

Are there restrictions regarding the reuse of data? What type of licenses need to be considered? How will researchers get access to your data? Will the data be publically available?

Preservation

What is your long-term preservation strategy? How much will long-term preservation cost? Where is the best place to preserve your data? How long does data need to be preserved? Will there be funding to assist in preserving data?

DMP Elements

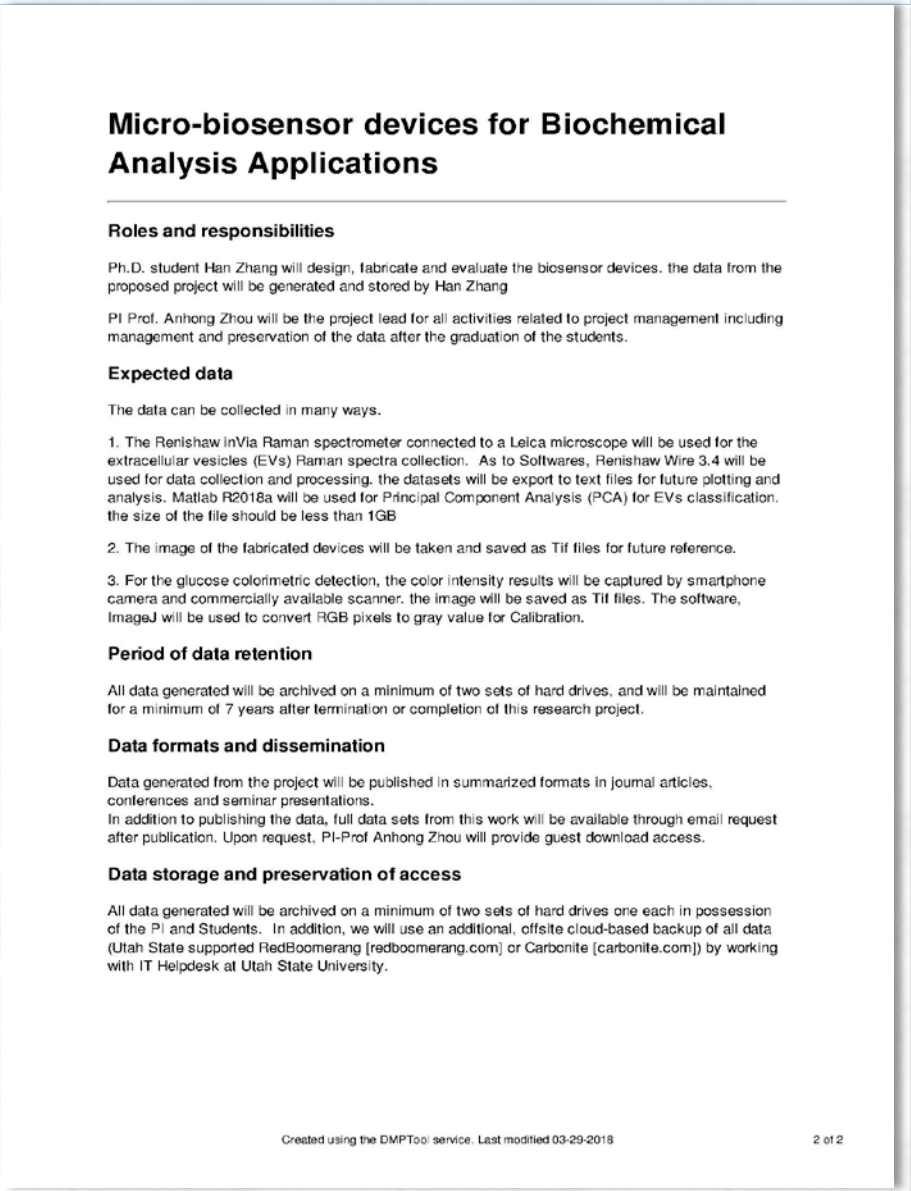
ROLES AND RESPONSIBILITIES
Who is the PI? Who else is collecting information? What is their roles in collecting data?

EXPECTED DATA
What sort of data is going to collect? How will data be collected? Is that data digital, analog, physical collections, biological? Is there special instrumentation?

PERIOD OF DATA RETENTION
How long will you retain the data? Will different formats of data be preserved differently?

DATA FORMATS AND DISSEMINATION
What are the specific data formats, media, and dissemination approaches that will be used to make data available to others? What are the policies for public access? How will data be described (E.g. metadata)?

DATA STORAGE AND PRESERVATION OF ACCESS
What are the physical and digital resources and facilities that will be used for the preservation of research data? What data repositories will be utilized? Where will you store the data for long term preservation? What are the costs associated with preservation and how will they be addressed?



Micro-biosensor devices for Biochemical Analysis Applications

Roles and responsibilities

Ph.D. student Han Zhang will design, fabricate and evaluate the biosensor devices. the data from the proposed project will be generated and stored by Han Zhang

PI Prof. Anhong Zhou will be the project lead for all activities related to project management including management and preservation of the data after the graduation of the students.

Expected data

The data can be collected in many ways.

1. The Renishaw inVia Raman spectrometer connected to a Leica microscope will be used for the extracellular vesicles (EVs) Raman spectra collection. As to Softwares, Renishaw Wire 3.4 will be used for data collection and processing. the datasets will be export to text files for future plotting and analysis. Matlab R2018a will be used for Principal Component Analysis (PCA) for EVs classification. the size of the file should be less than 1GB

2. The image of the fabricated devices will be taken and saved as Tif files for future reference.

3. For the glucose colorimetric detection, the color intensity results will be captured by smartphone camera and commercially available scanner. the image will be saved as Tif files. The software, ImageJ will be used to convert RGB pixels to gray value for Calibration.

Period of data retention

All data generated will be archived on a minimum of two sets of hard drives, and will be maintained for a minimum of 7 years after termination or completion of this research project.

Data formats and dissemination

Data generated from the project will be published in summarized formats in journal articles, conferences and seminar presentations.
In addition to publishing the data, full data sets from this work will be available through email request after publication. Upon request, PI-Prof Anhong Zhou will provide guest download access.

Data storage and preservation of access

All data generated will be archived on a minimum of two sets of hard drives one each in possession of the PI and Students. In addition, we will use an additional, offsite cloud-based backup of all data (Utah State supported RedBoomerang [redboomerang.com] or Carbonite [carbonite.com]) by working with IT Helpdesk at Utah State University.

Created using the DMPTool service. Last modified 03-29-2018 2 of 2

https://dmptool.org/plan_export/32453.pdf

DMPTool

Build your Data Management Plan

https://dmptool.org/plans/36933/edit?phase_id=409

Learn | David Luftig | English (US)

DMPTool
Build your Data Management Plan

University of Dayton

California Digital Library helpdesk

My Dashboard Create plan

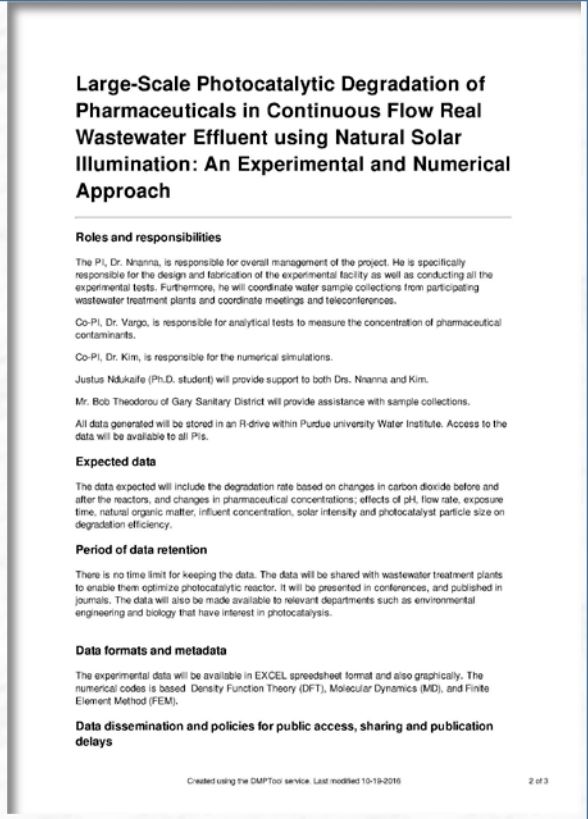
TEST

Project Details | Plan overview | Write Plan | Share | Download

expand all | collapse all 0/5 answered

+ Roles and responsibilities (0 / 1)
+ Expected data (0 / 1)
+ Period of data retention (0 / 1)
+ Data formats and dissemination (0 / 1)
+ Data storage and preservation of access (0 / 1)

https://dmptool.org/plans



Large-Scale Photocatalytic Degradation of Pharmaceuticals in Continuous Flow Real Wastewater Effluent using Natural Solar Illumination: An Experimental and Numerical Approach

Roles and responsibilities

The PI, Dr. Nnanna, is responsible for overall management of the project. He is specifically responsible for the design and fabrication of the experimental facility as well as conducting all the experimental tests. Furthermore, he will coordinate water sample collections from participating wastewater treatment plants and coordinate meetings and teleconferences.

Co-PI, Dr. Vargo, is responsible for analytical tests to measure the concentration of pharmaceutical contaminants.

Co-PI, Dr. Kim, is responsible for the numerical simulations.

Justus Ndukaife (Ph.D. student) will provide support to both Drs. Nnanna and Kim.

Mr. Bob Theodorou of Gary Sanitary District will provide assistance with sample collections.

All data generated will be stored in an R-drive within Purdue university Water Institute. Access to the data will be available to all PIs.

Expected data

The data expected will include the degradation rate based on changes in carbon dioxide before and after the reactors, and changes in pharmaceutical concentrations; effects of pH, flow rate, exposure time, natural organic matter, influent concentration, solar intensity and photocatalyst particle size on degradation efficiency.

Period of data retention

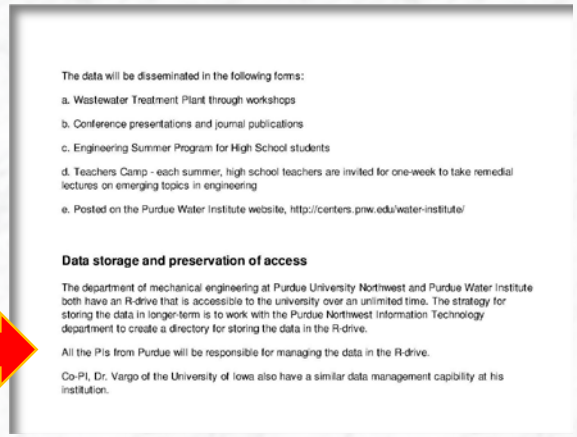
There is no time limit for keeping the data. The data will be shared with wastewater treatment plants to enable them optimize photocatalytic reactor. It will be presented in conferences, and published in journals. The data will also be made available to relevant departments such as environmental engineering and biology that have interest in photocatalysis.

Data formats and metadata

The experimental data will be available in EXCEL spreadsheet format and also graphically. The numerical codes is based Density Function Theory (DFT), Molecular Dynamics (MD), and Finite Element Method (FEM).

Data dissemination and policies for public access, sharing and publication delays

Created using the DMPTool service. Last modified 10-19-2016 2 of 3



The data will be disseminated in the following forms:

a. Wastewater Treatment Plant through workshops

b. Conference presentations and journal publications

c. Engineering Summer Program for High School students

d. Teachers Camp - each summer, high school teachers are invited for one-week to take remedial lectures on emerging topics in engineering

e. Posted on the Purdue Water Institute website, http://centers.pnw.edu/water-institute/

Data storage and preservation of access

The department of mechanical engineering at Purdue University Northwest and Purdue Water Institute both have an R-drive that is accessible to the university over an unlimited time. The strategy for storing the data in longer-term is to work with the Purdue Northwest Information Technology department to create a directory for storing the data in the R-drive.

All the PIs from Purdue will be responsible for managing the data in the R-drive.

Co-PI, Dr. Vargo of the University of Iowa also have a similar data management capibility at his institution.

https://dmptool.org/plan_export/23790.pdf

How the Roesch Library Can Assist You with Your DMP (Summary)

RESEARCH CONSULATIONS

- Conduct personalized interviews and consultations where we consider what sorts of data you will be collecting and help develop a management plan. Ideally this happens at the beginning of the research process.

WRITING DMPs

- We will sit down with you to find the DMP template that is most suitable for your discipline. We can than go through, element by element to fill create a DMP that is up to federal specifications.

MANAGING AND YOUR DATA

- We can examine how to manage your data short and long term in ways that meet federal requirements. We can also examine depositing digital data in federally approved data repositories. Additionally, we can examine issues of ownership and copyright.

RECOGNIZING WHERE TO PUBLISH YOUR RESEARCH

- We can assist in understanding issues of journal metrics, licenses, and open access.

Helpful Resources:

Roesch Library Data Management Research Guide

http://libguides.udayton.edu/datamanagement

Schedule a Data or Research Consultation

https://libcal.udayton.edu/appointments?u=47374

Direct email: dluftig1@udayton.edu

DMPTool

http://www.dmptool.org